Introduction to Computer Vision, 2012

Final Exam

- Answer 4 of the following 5 questions.
- You may write your answers either in Hebrew or in English.
- Length of exam: 3 hours.

Good luck!

Q1:

$f(x, y)$ is a continuous function whose maximal frequency in the x-direction and in the y-direction is Ω.

(a) $f(x, y)$ is discretely sampled at uniform distances. What is the minimal required sampling frequency in the x-direction and in the y-direction in order not to loose any information? (i.e., so that the continuous $f(x, y)$ can be reconstructed from its discrete samples.)

(b) The following operations are performed on the continuous function $f(x, y)$. For each case, explain what is the minimal required sampling frequency in the x-direction and in the y-direction in order not to loose any information:

(1) Shrinking of f by a factor of 3 in the x-direction, and stretching it by a factor of 2 in the y-direction:
$g(x, y) = f(3x, y/2)$.

(2) Shifting f:
$g(x, y) = f(x + 2, y - 3)$.

(3) Convolving f with another function h:
$g(x, y) = f(x, y) * h(x, y), \quad$ where $h(x, y) = cos(2\pi x)$.

Q2:

A moving camera is taking pictures of three scenes: a 3D scene (Scene 1), a planar scene (Scene 2), and a scene that is very distant from the camera (Scene 3). For each one of those scenes and each one of the following camera motions, explain whether the essential matrix can be computed. If not explain why. If it can be computed, describe the geometric arrangement of the epipolar lines. Provide a separate answer for each one of the above scenes and each camera motion below.

(Motion 1) A camera translating parallel to the image plane.

(Motion 2) A camera translating forward.

(Motion 3) A camera rotating around its optical axis.

Q3:

Two **stationary** video cameras record a 3-dimensional scene and a man moving within the scene. The cameras are attached to each other so that the distance between their focal centers is negligible. The cameras may have different zooms and different orientations, but both view the walking man. Assume that the cameras are temporally synchronized (i.e. their frames are taken at the same time)

(a) What is the geometric relation between the images obtained by the two cameras at each time?

(b) What is the minimal number of point correspondence required in order to recover this relation?

(c) Suppose the cameras are taking pictures in complete darkness, and suppose the man walks with a flashlight on his head (so that each camera sees the flashlight as a point). Under which of the following conditions can the transformation that relates the two images be recovered? Explain.

(1) The man walks on the sidewalk in a straight line.

(2) The man walks along the sidewalk and then turns and crosses the road.

(3) The man jumps up and down repeatedly in one place.

Q4:

A camera is moving while recording images. $I(x, y, t)$ is the image taken at time t. Let the spatial derivatives of the image intensity be denoted by I_x, I_y and the temporal derivative by I_t. We assume the Brightness Constancy Constraint at each pixel, whose linearized form is:

$$I_x u + I_y v + I_t = 0$$

where (u, v) is the velocity (displacement) of pixel (x, y) from time t to $t + 1$. Lucas and Kanade further assumed that within small image patches (e.g., 5×5, or 7×7), the displacement vector (u, v) is approximately constant. Therefore, for each pixel, they solve for the displacement (u, v) which minimizes the following error function within a small patch W surrounding that pixel:

$$Err(u, v) = \Sigma_{(x_i, y_i) \in W} \ (I_{x_i} u + I_{y_i} v + I_{t_i})^2. \tag{1}$$

(a) Derive the set of 2 linear equations in the 2 unknowns u and v that minimize the error function in Eq. (1).

(b) When will there be a unique solution, and when will there be degenerate cases (cases where the solution is not unique)? Explain both *algebraically*, as well *physically* (i.e., what is the underlying image structure within the image patch W that would give rise to a unique solution, or to singular cases).

Q5:

(a) Explain what is a HoG descriptor.

(b) Explain what is an SVM classifier.

(c) Describe briefly how you would use HoG and SVM to construct a classifier for car images.